JOURNAL OF MULTIVARIATE ANALYSIS 36, 44–55 (1991)

On Estimation of a Matrix of Normal Means with Unknown Covariance Matrix

YOSHIHIKO KONNO*

University of Tsukuba, Ibaraki 305, Japan

Communicated by the Editors

Let X be an $m \times p$ matrix normally distributed with matrix of means B and covariance matrix $I_m \otimes \Sigma$, where Σ is a $p \times p$ unknown positive definite matrix. This paper studies the estimation of B relative to the invariant loss function $\operatorname{tr} \Sigma^{-1}(\hat{B}-B)'(\hat{B}-B)$. New classes of invariant minimax estimators are proposed for the case $p > m+1$, which are multivariate extensions of the estimators of Stein and Baranchik. The method involves the unbiased estimation of the risk of an invariant estimator which depends on the eigenstructure of the usual $F = XS^{-1}X'$ matrix, where S: $p \times p$ follows a Wishart matrix with n degrees of freedom and mean $n\Sigma$.

1. INTRODUCTION

Assume that

$$\begin{aligned} &X;\, m \times p \sim N(B, I_m \otimes \Sigma) \\ &S;\, p \times p \sim W_p(\Sigma, n) \\ &X \text{ and } S \text{ are independent} \\ &B \text{ and } \Sigma \text{ are unknown.} \end{aligned} \tag{1.1}$$

Based on (X, S), we consider the problem of estimating B when the loss function is given by

$$L(\hat{B}, (B, \Sigma)) = \operatorname{tr} \Sigma^{-1}(\hat{B}-B)'(\hat{B}-B). \tag{1.2}$$

When $m=1$, James and Stein [2] obtained the estimator which dominates the usual estimator X for $p \geqslant 3$ relative to the loss function (1.2). Further-

Received October 24, 1989; revised March 30, 1990.

AMS 1980 subject classifications: primary 62F10; secondary 62C20.

Key words and phrases: minimax estimation, Stein estimator, Baranchik-type estimator.

* Present address: Ishinomaki Senshu University, Minamisakai, Ishinomaki-shi, Miyagi 305, Japan.

0047-259X/91 $3.00

more, Lin and Tsai [3] derived a family of minimax estimators. Our concern is how these results may be extended naturally to the multivariate case. Bilodeau and Kariya [1] treated the model (1.1) as the problem of estimating the coefficient matrix in a MANOVA model. Although they derived Efron-Morris-type estimator for $m > p+1$, they did not extensively exhibit the estimators for $p > m$. Here we extend the minimax estimators for $p \geqslant 3$ and $m = 1$ to the case for $p > m+1$. In Section 2 of this paper, we examine the estimator of the form $(I_m + 2RH(L)R')X$, where $L = \operatorname{diag}(l_1, ..., l_m)$ ($l_1 \geqslant \cdots \geqslant l_m \geqslant 0$ are eigenvalues of $XS^{-1}X'$ so that $XS^{-1}X' = RLR'$ and $RR' = R'R = I_m$) and $H(L) = \operatorname{diag}(\partial h(L)/\partial l_1, ..., \partial h(L)/\partial l_m)$ ($h(L)$ is a function from L to $[0, +\infty)$). This class of estimators are invariant under the transformation

$$X \to XA \qquad \text{and} \qquad S \to A'SA, \tag{1.3}$$

where A is a $p \times p$ nonsingular matrix. We obtain a condition under which $X + 2RH(L)R'X$ is minimax relative to the loss (1.2) by using the unbiased estimate of the risk (due to Bilodeau and Kariya [1]) of the estimator $X + G(X, S)$, where $G(X, S)$ is a $m \times p$ matrix function of X and S. The condition obtained here is similar to that due to Stein [7] for the case $\Sigma = I_p$ and facilitates an extensive search for superior alternatives to the commonly used estimator X. In Section 3, such search is undertaken. We specify the function $h(L)$ and obtain conditions under which the corresponding estimator is minimax. One of the improved estimators is

$$\left\{ I_m - \frac{p-m-1}{n+2m-p+1} (XS^{-1}X')^{-1} - \frac{m^2+m-2}{n-p+3} I_m / \operatorname{tr}(XS^{-1}X') \right\} X,$$

which is reduced to the Stein estimator for $m = 1$.

2. Unbiased Risk Estimate for Invariant Estimators

Let $F = XS^{-1}X'$. Let $g(F)$ be a function from F to $[0, +\infty)$. When $m = 1$, Haff [6] considered the estimators of the form

$$\hat{B} = (1 + 2g'(F))X, \tag{2.1}$$

where $g'(F) = dg(F)/dF$ and obtained the unbiased estimate of the risk for (2.1)

$$p + 4[pg'(F) + 2Fg''(F) - 4F^2 g'(F) g''(F) + (n-p-1) F\{g'(F)\}^2]. \tag{2.2}$$

A multivariate generalization of (2.1) and (2.2) may be defined in terms of the following differential operator: Let D_F be a differential operator (an $m \times m$ matrix) whose (i, j) element is given by $((1+\delta_{ij})/2))\,\partial/\partial F_{ij}$ for Kronecker's delta δ_{ij} and $F=(F_{ij})$. For a scalar function $g(F)$ and an $m \times q$ matrix $T=(t_{ij})$, operations by D_F are likewise defined;

$$D_F\, g(F);\, m \times m = \left(\frac{1+\delta_{ij}}{2} \frac{\partial g(F)}{\partial F_{ij}} \right),$$

and

$$D_F T;\, m \times q = \left(\sum_{k=1}^{m} \frac{1+\delta_{ik}}{2} \frac{\partial t_{kj}}{\partial F_{ik}} \right). \tag{2.3}$$

Then we consider estimators of the form

$$\hat{B} = (I_m + 2D_F\, g(F))\, X \tag{2.4}$$

which are invariant under the transformation (1.3).

Let $F=RLR'$ in which $RR'=R'R=I_m$, R is an $m \times m$ orthogonal matrix, and $L=\mathrm{diag}(l_1, ..., l_m)$ so that $l_1 \geqslant \cdots \geqslant l_m \geqslant 0$ are eigenvalues of F. If $p > m+1$, we can define another generalization of (2.1) given by

$$\hat{B} = (I_m + 2RH(L)\, R')\, X \tag{2.5}$$

where $H(L) = \mathrm{diag}(\partial h(L)/\partial l_1, ..., \partial h(L)/\partial l_m)$ and $\partial h(L)/\partial l_k$ $(k=1, ..., m)$ is absolutely continuous function from L to $[0, \infty)$. The estimator given by (2.4) includes those by (2.5) and both estimators coincide when $m=1$. First we shall derive the unbiased risk estimate of (2.4), which is a multivariate extension of (2.2). From this, we shall derive the unbiased estimate of risk for the estimators (2.5). For this end, we need the following notations and lemma due to Bilodeau and Kariya [1]. Let $\nabla_X;\, m \times p = (\partial/\partial X_{ij})$. Let D_S be a differential operator (a $p \times p$ matrix) whose (i, j) element is given by

$$((1+\delta_{ij})/2)(\partial/\partial S_{ij}).$$

Operations of D_S on a $p \times q$ matrix $T(S)$ are defined similarly in (2.3).

LEMMA 2.1 (Bilodeau and Kariya [1]). *Assume that $G(X, S)$ is an $m \times p$ matrix whose elements are absolutely continuous functions of X and S such that*

$$EG_{ij}^2 < \infty, \qquad E\,|\partial G_{ij}/\partial X_{kl}| < \infty \qquad \text{and} \qquad E(\partial G_{ij}/\partial S_{kl})^2 < \infty$$

and that the conditions in Theorem 2.1 (*Haff* [4]) *are satisfied. Then the unbiased estimate of the risk of the estimators* $X+G(X,S)$ *is given by*

$$\hat{R}=pm+\operatorname{tr}[2\nabla_X^t G(X,S)+2D_S G^t(X,S)\,G(X,S)$$
$$+(n-p-1)\,G^t(X,S)\,G(X,S)\,S^{-1}].$$

Furthermore, we record calculus lemma.

LEMMA 2.2. *Assume that* Q *and* T *are* $m\times q$ *matrix functions of* F *and that their derivatives exist as needed. Then we have*

(i) $\nabla_X^t Q=2S^{-1}X^t D_F Q$, *and if* $q=m$,

(ii) $\operatorname{tr}\nabla_X^t QX=p\operatorname{tr}Q+\operatorname{tr}X\nabla_X^t Q^t$,

(iii) $D_F Q T^t=(D_F Q)\,T^t+\{Q^t D_F^t\}^t\,T^t$,

(iv) $\operatorname{tr}(Q^t D_F^t)^t\,T^t=\operatorname{tr}Q^t D_F T$.

Remark 2.1. The familiar law for transposing products is not applicable in (iii). The product $Q^t D_F^t$ is computed, then the transpose is taken.

Proof. (i) The (i,j) element of $\nabla_X^t Q$ is equal to

$$\sum_{s_1=1}^{m}\sum_{s_2\leqslant s_3}\frac{\partial q_{s_1 j}}{\partial F_{s_2 s_3}}\cdot\frac{\partial F_{s_2 s_3}}{\partial X_{s_1 i}}=\sum_{s_1,s_2,s_3=1}^{m}\frac{1+\delta_{s_2 s_3}}{2}\cdot\frac{\partial q_{s_1 j}}{\partial F_{s_2 s_3}}\cdot\frac{\partial F_{s_2 s_3}}{\partial X_{s_1 i}},\tag{2.6}$$

where $Q=(q_{ij})$. From $F=XS^{-1}X^t$ and chain rule,

$$\frac{\partial F_{s_2 s_3}}{\partial X_{s_1 i}}=\sum_{s_4=1}^{p}\{\delta_{s_1 s_2}S^{is_4}X_{s_3 s_4}+\delta_{s_1 s_3}S^{is_4}X_{s_2 s_4}\},\tag{2.7}$$

where $S^{-1}=(S^{ij})$. Putting (2.7) into (2.6) and using the symmetry of F, we get the desired result.

(ii) The proof follows from the straightforward calculation.

(iii) See Haff [5].

(iv) The proof follows from the straightforward calculation.

THEOREM 2.1. *If conditions of Lemma* 2.1, *where* $G(X,S)=2D_F\,g(F)\,X$ *are satisfied, the unbiased estimate of the risk for the estimators* (2.4) *is given by*

$$pm+4\operatorname{tr}[pD_F\,g(F)+2FD_F^2\,g(F)+(n+2m-p+1)\,F\{D_F\,g(F)\}^2$$
$$-4(D_F\,g(F))\,FD_F\{FD_F\,g(F)\}].\tag{2.8}$$

Proof. From Lemma 2.1, the unbiased estimate of the risk for $X+2D_F g(F)X$ can be expressed as

$$\begin{aligned}&4\,\mathrm{tr}\,\nabla'_X D_F g(F)X+8\,\mathrm{tr}\,D_S\{X'(D_F g(F))^2X\}\\&\quad+4(n-p-1)\,\mathrm{tr}\,F(D_F g(F))^2+pm.\end{aligned}\tag{2.9}$$

We shall calculate the above term by term. From (i) and (ii) of Lemma 2.2 and symmetry of $D_F g(F)$, it follows that the first term in (2.9) becomes

$$4p\,\mathrm{tr}\,D_F g(F)+4\,\mathrm{tr}\,X\nabla'_X D_F g(F)=4p\,\mathrm{tr}\,D_F g(F)+8\,\mathrm{tr}\,FD_F^2 g(F).\tag{2.10}$$

Next, applying (iii) of Lemma 2.2 with D_S instead of D_F to the second term in (2.9) provides that

$$\begin{aligned}&8\,\mathrm{tr}\,D_S\{X'D_F g(F)\}(D_F g(F))X+8\,\mathrm{tr}[\{X'D_F g(F)\}'D'_S]'D_F g(F)X\\&\quad=16\,\mathrm{tr}(D_F g(F))XD_S\{X'D_F g(F)\}.\end{aligned}\tag{2.11}$$

The equality holds follows from (iv) of Lemma 2.2. Using ordinary chain rule and (iii) of Lemma 2.2, straightforward calculation shows that

$$XD_S\{X'D_F g(F)\}=-FD_F(FD_F g(F))+((m+1)/2)\,FD_F g(F).\tag{2.12}$$

See the Appendix for the verification of (2.12). Hence, putting (2.12) into (2.11), it is seen that the second term in (2.9) can be rewritten as

$$-16\,\mathrm{tr}(D_F g(F))\cdot FD_F\{FD_F g(F)\}+8(m+1)\,\mathrm{tr}\,F\{D_F g(F)\}^2.\tag{2.13}$$

Finally, combining (2.10) and (2.13) with (2.9) leads to the desired result.

In the following we assume that $m+1<p$. Here, we record formulas for D_F and its element operating on eigenstructure from Haff [6] which are useful in deriving the unbiased risk estimate of (2.5).

LEMMA 2.3. *Let $R=(R_1, R_2, ..., R_m)$ (so R_i is the eigenvector corresponding to l_i) and $\phi(L)=\mathrm{diag}(\phi_1(L), ..., \phi_m(L))$, where $\phi_k(L)$ $(k=1, ..., m)$ is a function from L to $[0, +\infty)$. Assuming that all relevant derivatives exist, we have*

(i) $D_F l_i=R_iR'_i$,

(ii) $D_F R_i=l_i^*R_i$, $l_i^*=\frac{1}{2}\sum_{t\neq i}1/(l_i-l_t)$, *and*

(iii) $D_F[R\phi(L)R']=R\phi^{(1)}(L)R'$, *where* $\phi^{(1)}(L)=\mathrm{diag}(\phi_1^{(1)}(L), ..., \phi_m^{(1)}(L))$ *and*

$$\phi_k^{(1)}(L)=\frac{1}{2}\sum_{t\neq k}(\phi_k(L)-\phi_t(L))/(l_k-l_t)+\partial\phi_k(L)/\partial l_k,\qquad k=1, ..., m.$$

Proof. See Haff [6] and Stein [6].

Remark 2.2. Set $H(L) = \mathrm{diag}(\partial h(L)/\partial l_1, ..., \partial h(L)/\partial l_m)$, where $h(L)$ is a function from L to $[0, +\infty)$. From (i) of Lemma 2.2 and (i) of Lemma 2.3, we can see that

$$2RH(L)\, R'X = 2D_F h(L)\, X = \nabla_X h(L)\, S$$

which follows that the estimator $[I_m + 2RH(L)\, R']\, X$ can be rewritten as

$$[I_m + 2D_F h(L)]\, X = X + \nabla_X h(L)\, S. \tag{2.14}$$

THEOREM 2.2. *The improvement of the risk of* (2.5) *or equivalently* (2.14) *over that of X becomes*

$$\begin{aligned} &E_{B,\Sigma}[\|X - B\|^2 - \|X + \nabla_X h(L)\, S - B\|^2] \\ &= -4E_{B,\Sigma}\left[\sum_{k=1}^{m} \left\{(p - m + 1)\, h_k(L) + 2l_k h_{kk}(L)\right.\right. \\ &\quad + 2\sum_{t>k} \frac{l_k h_k(L) - l_t h_t(L)}{l_k - l_t} + (n + 2m - p - 3)\, l_k h_k^2(L) \\ &\quad \left.\left. - 4l_k^2 h_k(L)\, h_{kk}(L) - 2\sum_{t>k} \frac{l_k^2 h_k^2(L) - l_t^2 h_t^2(L)}{l_k - l_t}\right\}\right], \end{aligned} \tag{2.15}$$

where $h_k(L) = \partial h(L)/\partial l_k$ *and* $h_{kk}(L) = \partial^2 h(L)/\partial l_k^2$.

Proof. Set $g(F) = h(L)$ in (2.8). Then we can see that the left side of (2.15) provides

$$\begin{aligned} &-4E_{B,\Sigma}[\mathrm{tr}\{pD_F h(L) + 2FD_F^2 h(L) + (n + 2m - p + 1)\, F\{D_F h(L)\}^2 \\ &\quad - 4(D_F h(L))\, FD_F\{FD_F h(L)\}\}]. \end{aligned} \tag{2.16}$$

Using (iii) of Lemma 2.3 it follows that

$$\mathrm{tr}\, D_F h(L) = \sum_{k=1}^{m} h_k(L)$$

and

$$\mathrm{tr}\, F\{D_F h(L)\}^2 = \sum_{k=1}^{m} l_k h_k^2(L). \tag{2.17}$$

Noting that $D_F^2 h(L) = D_F RH(L) R'$, where $H(L) = \mathrm{diag}(h_1(L), ..., h_m(L))$ and, applying (iii) of Lemma 2.3, where $\phi(L) = H(L)$, we can find that

$$\begin{aligned}\mathrm{tr}\, FD_F^2 h(L) &= \sum_{k=1}^{m} \left[l_k \left\{ h_{kk}(L) + \frac{1}{2} \sum_{t \neq k} (h_k(L) - h_t(L))/(l_k - l_t) \right\} \right] \\ &= \sum_{k=1}^{m} \left[l_k (h_{kk}(L) - ((m-1)/2)\, h_k(L)) \right. \\ &\quad \left. + \sum_{t > k} (l_k h_k(L) - l_t h_t(L))/(l_k - l_t) \right]. \end{aligned} \tag{2.18}$$

Similarly, we can obtain that

$$\begin{aligned} &\mathrm{tr}(D_F h(L))\, FD_F \{FD_F h(L)\} \\ &= \sum_{k=1}^{m} \left[l_k h_k(L) \left\{ h_k(L) + l_k h_{kk}(L) \right. \right. \\ &\quad \left. \left. + \frac{1}{2} \sum_{t \neq k} \{ l_k h_k(L) - l_t h_t(L) \} / (l_k - l_t) \right\} \right] \\ &= \sum_{k=1}^{m} \left[l_k h_k^2(L) + l_k^2 h_k(L)\, h_{kk}(L) \right. \\ &\quad \left. + \frac{1}{2} \sum_{t > k} \{ l_k^2 h_k^2(L) - l_t^2 h_t^2(L) \} / (l_k - l_t) \right]. \end{aligned} \tag{2.19}$$

Finally, putting (2.17), (2.18), and (2.19) into (2.16), we obtain the desired result.

3. Minimax Estimators

In this section, some specifications of $h(L)$ in (2.5) or (2.14) are made and sufficient conditions under which the corresponding estimator $\hat{B}(X, S)$ is minimax relative to the loss (1.2) are obtained.

Using Theorem 2.2, we get a multivariate generalization of Stein estimator.

Theorem 3.1. *When $p > m + 1$, the estimator*

$$\hat{B}^{(1)}(X, S) = [I_m - a(XS^{-1}X')^{-1} - bI_m/\mathrm{tr}\, XS^{-1}X']\, X \tag{3.1}$$

is minimax relative to the loss (1.2) *if $a = (p - m - 1)/(n + 2m - p + 1)$ and $b = (m^2 + m - 2)/(n - p + 3)$.*

Proof. Let

$$h^{(1)}(L) = -\frac{1}{2}\left\{\log \prod_{k=1}^{m} l_k^a + \log\left(\sum_{k=1}^{m} l_k\right)^b\right\}, \tag{3.2}$$

where a and b are nonnegative constants. Using (i) of Lemma 2.2 the estimator (2.14) with (3.2) becomes the estimator of the form (3.1). Set $h_k^{(1)} = \partial h_k^{(1)}/\partial l_k$ and $h_{kk}^{(1)} = \partial^2 h_k^{(1)}/\partial l_k^2$, $k = 1, ..., m$. We observe

$$h_k^{(1)} = -\frac{1}{2}\left(\frac{a}{l_k} + \frac{b}{u}\right),$$

$$h_{kk}^{(1)} = \frac{1}{2}\left(\frac{a}{l_k^2} + \frac{b}{u^2}\right),$$

$$2 \sum_{t>k} \frac{l_k h_k^{(1)} - l_t h_t^{(1)}}{l_k - l_t} = -\frac{m-1}{2}\frac{b}{u},$$

$$2 \sum_{t>k} \frac{l_k^2 (h_k^{(1)})^2 - l_t^2 (h_t^{(1)})^2}{l_k - l_t} = -\frac{(m-1)\,ab}{2u} - \frac{(m-1)\,b^2}{2mu},$$

where $u = \sum_{k=1}^{m} l_k$. Noting that $\sum_{k=1}^{m} m l_k^2/u^2 \leqslant m$ and putting the above equations into (2.15), we obtain

$$\begin{aligned} \Delta &= E\{[\|X - B\|^2 - \|\hat{B}^{(1)} - B\|^2]\} \\ &\geqslant -E\left\{\sum_{k=1}^{m}\left[\frac{1}{l_k}((n+2m-p+1)\,a^2 - 2(p-m-1)\,a)\right.\right. \\ &\quad \left.\left. + \frac{1}{mu}\{(n-p+3)\,b^2 + 2(m(n+m-p)+2)\,ab - 2(mp-2)\,b\}\right]\right\}. \end{aligned} \tag{3.3}$$

The first term of the right side of (3.3) is minimized when $a = (p-m-1)/(n+2m-p+1)$, in which the term is negative. For $a = (p-m-1)/(n+2m-p+1)$, the second term is bounded above

$$(n-p+3)\,b^2 - 2(m^2+m-2)\,b.$$

It is minimized when $b = (m^2+m-2)/(n-p+3)$ in which the term is negative. This completes the proof.

Next, we give another generalization of Stein estimator.

THEOREM 3.2. *When $p > m+1$, the estimator*

$$\hat{B}^{(2)}(X, S) = [I_m - RH^{(2)}(L)\,R']\,X, \tag{3.4}$$

where $H^{(2)}(L) = \mathrm{diag}(d_1/l_1, ..., d_m/l_m)$ *and* $d_k \geqslant \cdots \geqslant d_m$ *are nonnegative constants, is minimax relative to the loss* (1.2) *if* $d_k = (p+m-1-2k)/(n-p+1+2k)$, $k = 1, ..., m$.

Proof. Let

$$h^{(2)}(L) = -\sum_{k=1}^{m} \frac{d_k}{2} \log l_k$$

in (2.14). Similar calculation to the proof of Theorem 3.1 shows that the risk difference between X and the estimator given by (3.4) is

$$\begin{aligned} \Delta &= E[\|X-B\|^2 - \|\hat{B}^{(2)} - B\|^2] \\ &= -E\left[\sum_{k=1}^{m} \left\{-2(p-m-1)\frac{d_k}{l_k} + (n+2m-p+1)\frac{d_k^2}{l_k}\right\}\right] \\ &\quad + E\left[\sum_{k=1}^{m} \sum_{t>k} \frac{4d_k + 2d_k^2 - (4d_t + 2d_t^2)}{l_k - l_t}\right]. \end{aligned} \tag{3.5}$$

Set $y_k = 4d_k + 2d_k^2$. Note that $y_1 \geqslant \cdots \geqslant y_m$. Then we obtain

$$\begin{aligned} \sum_{k=1}^{m} \sum_{t>k} \frac{y_k - y_t}{l_k - l_t} &= \sum_{k=1}^{m} \frac{1}{l_k} \sum_{t>k} \frac{l_k}{l_k - l_t}(y_k - y_t) \\ &\geqslant \sum_{k=1}^{m} \frac{1}{l_k} \sum_{t>k} (y_k - y_t) \\ &= \sum_{k=1}^{m} \frac{(m-k)\, y_k - \sum_{t>k} y_t}{l_k} \geqslant 0, \end{aligned} \tag{3.6}$$

since $l_k/(l_k - l_t) > 1$ for $t > k$. From above, (3.5) is bounded below

$$-E\left[\sum_{k=1}^{m} \frac{1}{l_k}\left\{(n-p+1+2k)\, d_k^2 - 2(p+m-1-2k)\, d_k + \sum_{t>k} (4d_t + 2d_t^2)\right\}\right].$$

Define

$$\begin{aligned} z_k(d_k) = {} & (n-p+1+2k)\, d_k^2 - 2(p+m-1-2k)\, d_k \\ & + \sum_{t>k} (4d_t + 2d_t^2), \qquad k = 1, ..., m, \end{aligned}$$

where $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_m$. It is sufficient to prove that $z_k(d_k)$ is negative when $d_k = (p+m-1-2k)/(n-p+1+2k)$ (say d_k^0), $k = 1, ..., m$. We shall fix $d_{k+k'}=d^0_{k+k'}$ $(k'=1, ..., m-k)$ to choose d_k. Then we can see that $z_k(d_k)$ is minimized when $d_k=d^0_k$ and that

$$z_k(d^0_k)<z_k(d^0_{k+1})=z_{k+1}(d^0_{k+1})< \cdots <z_m(d^0_m)<0,$$

since $d^0_1> \cdots >d^0_m$. This completes the proof.

From the proof above, it turns out that $\hat{B}^{(2)}$ is better than $X-\{(p-m-1)/(n+2m-p+1)\}F^{-1}X$ when $m \geqslant 2$, although both estimators reduce to the usual Stein estimator when $m=1$.

The next theorem is a generalization of Lin and Tsai's result [2].

THEOREM 3.3. *Let $\gamma_k(t)$ $(k=1, ..., m)$ be functions satisfying*

(i) $0 \leqslant \gamma_k(t) \leqslant 2(p-m-1)/(n+2m-p+1)$, $k=1, ..., m$,

(ii) $\gamma_k(t)$ *is nondecreasing,*

(iii) $\gamma_1(t) \geqslant \gamma_2(t) \geqslant \cdots \geqslant \gamma_m(t)$ *for* $\forall t \geqslant 0$.

Then the estimator $\hat{B}^{(\gamma)}(X, S)=[I_m-RH^{(\gamma)}(L)\,R']\,X$, *where* $\gamma=(\gamma_1, ..., \gamma_m)$ *and*

$$H^{(\gamma)}(L)=\operatorname{diag}\left(\frac{\gamma_1(l_1)}{l_1}, \frac{\gamma_2(l_2)}{l_2}, ..., \frac{\gamma_m(l_m)}{l_m}\right),$$

is minimax for $p>m+1$.

Proof. Similar to the proof of Theorem 2 in Zheng [8], first we suppose that $\gamma_k(t)$, $k=1, ..., m$, are absolutely continuous and have bounded derivatives on $[0, \infty)$. We shall use the notation $\hat{B}(\gamma)$ instead of $\hat{B}^{(\gamma)}(X, S)$ for convenience. Using Theorem 2.2, we obtain that

$$\begin{aligned}
\Delta &= E[\|X-B\|^2]-E[\|\hat{B}(\gamma)-B\|^2] \\
&= -E\left[\sum_{k=1}^{m}\left\{\frac{(n+2m-p+1)\,\gamma_k(l_k)}{l_k}\left(\gamma_k(l_k)-\frac{2(p-m-1)}{n+2m-p+1}\right)\right\}\right] \\
&\quad +E\left[\sum_{k=1}^{m}\left\{4(1+\gamma_k(l_k))\frac{\partial \gamma_k(l_k)}{\partial l_k}+\sum_{t>k}\left\{1+\frac{1}{2}(\gamma_k(l_k)+\gamma_t(l_t))\right\}\right.\right. \\
&\quad \left.\left.\times\left\{\frac{\gamma_k(l_k)-\gamma_t(l_t)}{l_k-l_t}\right\}\right\}\right]. \qquad (3.7)
\end{aligned}$$

From the above and the condition on γ_k, we find that $\Delta \geqslant 0$, which follows that $\hat{B}(\gamma)$ is minimax.

Suppose now that $\gamma_k(t)$'s, $k=1, ..., m$, are general functions satisfying the conditions of Theorem. Let $\gamma^{(i)}=(\gamma^{(i)}_1, ..., \gamma^{(i)}_m)$, where $\gamma^{(i)}_k$'s, $k=1, ..., m$, are functions which are absolutely continuous and have bounded derivatives such that $\gamma_k^{(i)}(t)$ converges to $\gamma_k(t)$ as $i \to +\infty$. Then the estimator $\hat{B}(\gamma^{(i)})$ converges to $\hat{B}(\gamma)$ (a.e.) as $i \to +\infty$. From (3.7), it follows that $\hat{B}(\gamma^{(i)})$ is minimax and $\|\hat{B}(\gamma^{(i)}) - B\|^2$ has the bounded expectation independent of i. So $\hat{B}(\gamma)$ is minimax. This completes the proof.

Appendix

The proof in the line of (2.12). Set $D_S = (d_S^{ij})$ and $D_F = (d_F^{ij})$. It can be seen that the (i, j) element of $XD_S\{X'D_F g(F)\}$ is

$$\begin{aligned}
&\sum_{t_1, t_2 = 1}^{p} \sum_{t_3 = 1}^{m} X_{it_1} d_S^{t_1 t_2} X_{t_3 t_2} d_F^{t_3 j} g(F) \\
&= \sum_{t_1, t_2 = 1}^{p} \sum_{t_3, u_1, u_4 = 1}^{m} X_{it_1} X_{t_3 t_2} (d_F^{u_1 u_4} d_F^{t_3 j} g(F))\, d_S^{t_1 t_2} F_{u_1 u_4} \\
&= \sum_{t_1, t_2, u_2, u_3 = 1}^{p} \sum_{t_2, u_1, u_4 = 1}^{m} X_{it_1} X_{t_3 t_2} X_{u_1 u_2} X_{u_4 u_3} (d_F^{u_1 u_4} d_F^{t_3 j} g(F))\, d_S^{t_1 t_2} S^{u_2 u_3}, \qquad (1)
\end{aligned}$$

where the last equality can be obtained by putting $F_{u_1 u_4} = \sum_{u_2, u_3 = 1}^{p} X_{u_1 u_2} S^{u_2 u_3} X_{u_4 u_3}$. Using $d_S^{t_1 t_2} S^{u_2 u_3} = -\frac{1}{2}(S^{u_2 t_1} S^{t_2 u_3} + S^{u_2 t_2} S^{t_1 u_3})$ (see Haff [4]), the formula (1) becomes $-F_{iu_1} F_{t_3 u_4} d_F^{u_1 u_4} d_F^{t_3 j} g(F)$, from which it follows that

$$\begin{aligned}
XD_S\{X'D_F g(F)\} &= -F(FD_F)' D_F g(F) \\
&= -FD_F(FD_F g(F)) + ((m+1)/2)\, FD_F g(F).
\end{aligned}$$

The last equality holds, from (iii) of Lemma 2.2 and $D_F F = ((m+1)/2)\, I_m$.

Acknowledgments

The author thanks Professor N. Sugiura and Dr. T. Kubokawa for their helpful comments and encouragement. He is also grateful to the referee for providing critical comments which led to a much improved version of the paper.

References

[1] Bilodeau, M., and Kariya, T. (1989). Minimax estimators in the normal MANOVA model. *J. Multivariate Anal.* **28** 260–270.

[2] James, W., and Stein, C. (1961). Estimation with quadratic loss. In *Proceedings, Fourth Berkeley Symp. Math. Statist. Probab.*, Vol. 1, pp. 361–379, Univ. of California Press, Berkeley.

[3] LIN, P.-E., AND TSAI, H.-L. (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. *Ann. Statist.* **1** 142–145.

[4] HAFF, L. R. (1979). An identity for the Wishart distribution with applications. *J. Multivariate Anal.* **9** 531–544.

[5] HAFF, L. R. (1981). Further identities for the Wishart distribution with applications in regression. *Canad. J. Statist.* **9** 215–224.

[6] HAFF, L. R. (1982). Solutions of the Euler–Lagrange equations for certain multivariate normal estimation problem. Unpublished manuscript.

[7] STEIN, C. (1973). Estimation of the mean of a multivariate normal distribution. In *Proceedings, Prague Symp. Asymptotic Statist.*, pp. 345–381.

[8] ZHENG, Z. (1986). On estimation of matrix of normal mean. *J. Multivariate Anal.* **18** 70–82.